

# Real-time implementation of MUSIC for wideband acoustic detection and tracking

Tien Pham and Manfai Fong

U.S. Army Research Laboratory  
Adelphi, MD 20783-1197

## ABSTRACT

The U.S. Army Research Laboratory (ARL) has been developing advanced acoustic array signal processing algorithms using small baseline arrays for detecting, performing direction finding, tracking, and classifying ground targets. In this paper, we discuss the wideband MUSIC algorithms and the real-time implementation of these algorithms in the ARL sensor testbed. Computational complexity issues and CPU platforms pertaining to the testbed are discussed. In addition, we present experimental results for multiple targets test runs showing the relative performance of the delay-sum and the incoherent wideband MUSIC algorithms versus ground truth.

**Key words:** acoustics, wideband, MUSIC, direction of arrival (DOA), high-resolution, real-time, ground vehicles

## 1. INTRODUCTION

This paper describes the current work on array signal processing algorithms that passively detect, estimate DOA, and track moving ground targets. We also describe the real-time implementation of these algorithms in the ARL sensor testbed [1]. In our application, we are motivated to use high resolution methods such as MUSIC for small baseline acoustic arrays because of the system requirements and, in general, the lack of spatial coherence between the sensing elements beyond several meters. High-resolution capability is needed for accurate and reliable detection and tracking of multiple targets and convoys.

Acoustic detection and tracking of ground vehicles in a battlefield environment is a challenging problem. The acoustic signature of ground vehicles are wideband and nonstationary. Most tracked and wheeled vehicles of interest have diesel engines, which exhibit a pronounced harmonic structure (i.e., sharp spectra) due to the engine rates and, to a lesser extent, the track slaps. However, most of the detectable spectral peaks are below 200 Hz except at very close range (<200 m). The atmospheric absorption attenuates the higher frequencies more rapidly at longer ranges. In addition, below 20 Hz, the acoustic signature is corrupted by wind noise [2]. Other vehicles of interest have turbine engines which exhibit broadband energy from less than 20 Hz to beyond 2 kHz. Some spectral harmonic peaks are noticeable, due to the track slaps, in some test runs depending on the test environment. The detectable frequency range of the turbine engine vehicles is dependent upon the range of the target [3]. Unlike the signatures of helicopters and small aircraft, acoustic signatures of ground vehicles are generally nonstationary and undergo severe fading during maneuvers. The combined effects of source, terrain, and propagation medium produce large signal variability, even at relatively close range [1].

The challenge is to develop real-time adaptive, robust, high-resolution, wideband algorithms that will improve the accuracy and stability of the DOA estimates for multiple target detection and tracking. In this application, a one-second update is considered real-time. Current efforts are focused on incoherent and coherent wideband MUSIC [1, 2, 3]. Validation is accomplished using experimental data and ground truth, and successful algorithms are being implemented in the sensor testbed. DOA results using a circular array of six sensors plus one sensor at the

---

Further author information

T. P. (Correspondence): Email: tpham@arl.mil; WWW: <http://www.arl.mil/acoustics>; Tel: (301)-394-4282; Fax: (301)394-2593  
M. F.: Email: mfong@arl.mil; WWW: <http://www.arl.mil/acoustics>; Tel: (301)-394-4282; Fax: (301)394-2593

array center, with a diameter of 8 ft are presented for wideband MUSIC. We discuss the algorithm performance, computational complexity, and real-time implementation issues.

## 2. SENSOR TESTBED

The sensor testbed is an evolving distributed platform used to prototype and evaluate new passive sensor technology in both hardware and software applications in real-time (see figure 1). Multiple-sensor arrays transmit detection, DOA, and classification information back to a central gateway on a second-by-second basis. The gateway performs data fusion by collating DOA estimates with a tracker and classification reports with a classifier, and transmits results back to the combat information processor (CIP) for display on a graphical user interface. The testbed is designed so that any type of sensor or array geometry can be used, as long as it adheres to a predetermined communications protocol. The gateway is designed so that any data fusion algorithm can be used, as long as it uses predetermined communication protocols [4]. The CIP remotely displays how well the sensor arrays detect vehicles, and how well the data fusion algorithm at the gateway is correlating the multiple-sensor arrays reports. Packet radios are used for communication links between the sensor arrays and the gateway, and between the gateway and the CIP.

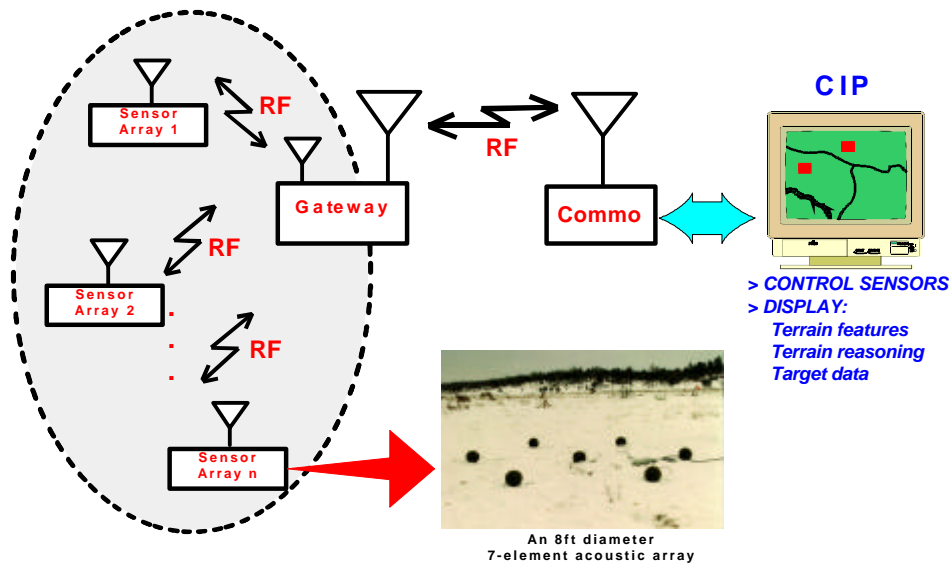


Figure 1. The ARL sensor testbed system overview.

The current sensor testbed deploys multiple arrays of 7 microphones (see figure 1), or arrays of 5 microphones with one 3-axis geophone [5]. The seismic sensor provides some unique features that complement well with acoustic features. Both acoustic and seismic features are being used to train a neural network classifier [6]. Crude direction finding (DF) can be accomplished using a 3-axis geophone if the velocity in the ground is known. Other low-cost passive sensors of interest that can be incorporated into the testbed are magnetic and electrostatic sensors.

## 3. WIDEBAND MUSIC ALGORITHMS

### 3.1 Background

Conventional delay-sum (DS) beamforming and narrowband MUSIC have been implemented at the sensor-array level. The DS beamformers perform well for a single-source environment; however, their ability to resolve closely spaced sources is limited by the small baselines of the arrays. Narrowband MUSIC produces sharp beampatterns, but requires the frequency bin to have high SNR. For multiple targets, it requires sophisticated peak picking and data association algorithms for successful implementation. In general, any narrowband DOA technique used for this application will not fully exploit the wideband nature of the acoustic sources. We seek to develop and implement wideband DOA algorithms for further processing gains. We want to exploit as much of the multi-

spectral content from the acoustic sources as possible, improve accuracy and stability of the DOA estimates, and produce higher resolution outputs.

### 3.2 Wideband Implementation of MUSIC

Figure 2 shows the flow diagram describing the implementation steps of incoherent and coherent wideband MUSIC algorithms. The front-end preprocessing steps are identical for both wideband algorithms. To overcome the nonstationary nature of the acoustic source, the data are segmented before processing into fixed blocks, and stationarity is assumed for each data block. We have found that the assumption of signal stationarity is reasonable for most instances, for intervals on the order of 1 second or less, with sampling rates of 1 kHz and 2 kHz. For each data block, the first step is to compute  $\hat{X}_i(\mathbf{w}), i=1,2,\dots,N$  ( $N$  is the number of elements in the sensor array), the fast Fourier transform (FFT) of the corresponding time series data  $\hat{x}_i(n)$ . The second step is to compute the average power spectrum of all sensors, and then adaptively select the  $M$  frequency components,  $\hat{X}(\mathbf{w}_k)=[\hat{X}_1(\mathbf{w}_k), \hat{X}_2(\mathbf{w}_k), \dots, \hat{X}_N(\mathbf{w}_k)]$  for  $k=1,2,\dots,M$ , for wideband processing. This can be performed in a variety of ways, from simple thresholding based on frequency bin signal-to-noise ratio (SNR), to more complex schemes, such as harmonic line association (HLA) [5,6]. The next step is to form the estimated narrowband spatial correlation matrix,  $\hat{R}_X(\mathbf{w}_k)$ , at each frequency  $\mathbf{w}_k$  for  $k=1,2,\dots,M$  over each data block.

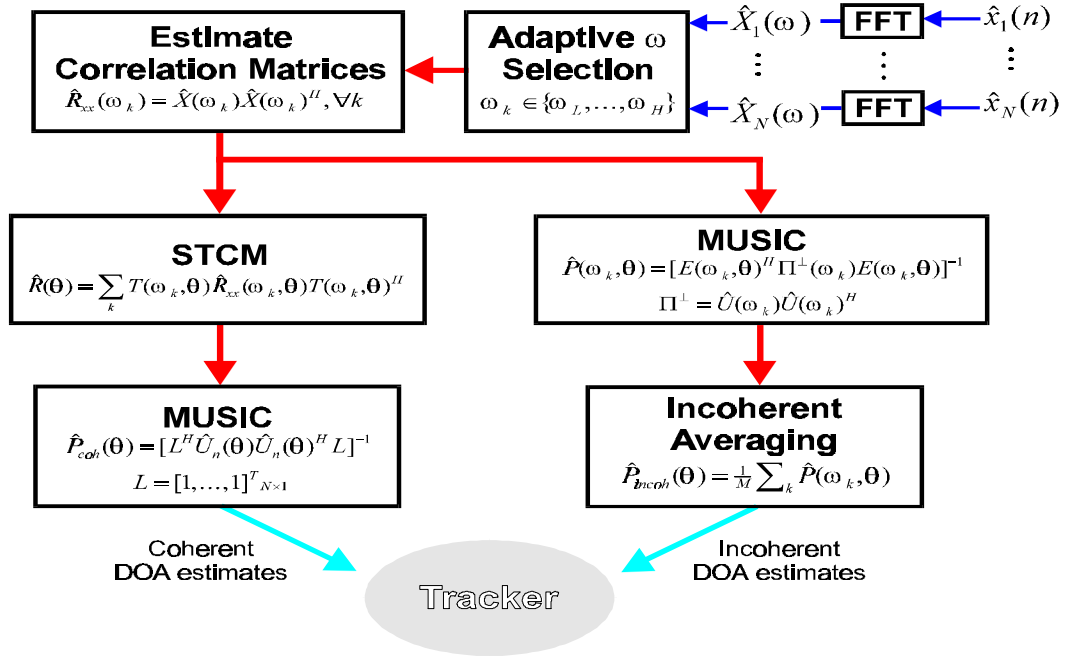


Figure 2. Implementation of the incoherent and coherent wideband MUSIC algorithms.

For incoherent wideband MUSIC, the narrowband MUSIC pseudospectrum or beampattern is computed from each  $\hat{R}_X(\mathbf{w}_k)$ , and then the beampatterns are incoherently averaged together to form a resulting beampattern,  $\hat{P}_{Incoh}(\mathbf{q})$ , given by

$$\hat{P}_{Incoh}(\mathbf{q}) = \sum_{k=1}^M [E(\mathbf{w}_k, \mathbf{q})^H \hat{U}_n(\mathbf{w}_k) \hat{U}_n(\mathbf{w}_k)^H E(\mathbf{w}_k, \mathbf{q})]^{-1}. \quad (1)$$

$E(\mathbf{w}_m, \mathbf{q})$  is the steering vector, defined as

$$E(\mathbf{w}_k, \mathbf{q}) = [e^{j\mathbf{w}_k \Delta t_1}, e^{j\mathbf{w}_k \Delta t_2}, \dots, e^{j\mathbf{w}_k \Delta t_M}]^T, \quad (2)$$

where  $\Delta t_i = d/c \sin \mathbf{f}_i$ ,  $\mathbf{f}_i = \mathbf{q} - \mathbf{a}_i$  is the angle relative to the normal for sensor  $i$ ,  $i=1, 2, \dots, N$ ,  $d$  is the radius of the circular array, and  $c$  is the speed of sound in air.  $\hat{U}_n(\mathbf{w}_k)$  is the noise subspace estimate at  $\mathbf{w}_k$  [2,3].

For coherent wideband processing,  $\hat{R}_X(\mathbf{w}_k)$ s are combined to form one focused or steered correlation matrix,  $\hat{R}(\mathbf{q})$ , at each look angle  $\mathbf{q}$ .  $\hat{R}(\mathbf{q})$  is defined as

$$\hat{R}(\mathbf{q}) = \sum_{k=1}^M T(\mathbf{w}_k, \mathbf{q}) \hat{R}_X(\mathbf{w}_k) T(\mathbf{w}_k, \mathbf{q})^H, \quad (3)$$

where  $T(\mathbf{w}_k, \mathbf{q})$  is the diagonal matrix of the steering vector  $E(\mathbf{w}_m, \mathbf{q})$ . This focusing step is based on the steered covariance matrix (STCM) method [2,3,8]. The next step is performing the MUSIC algorithm on  $\hat{R}(\mathbf{q})$  to yield a coherent wideband beampattern,  $\hat{P}_{\text{Coh}}(\mathbf{q})$ , given by

$$\hat{P}_{\text{Coh}}(\mathbf{q}) = [L^H \hat{U}_n(\mathbf{q}) \hat{U}_n(\mathbf{q})^H L]^{-1}. \quad (4)$$

$\hat{U}_n(\mathbf{q})$  is unitary noise subspace estimate at look angle  $\mathbf{q}$ , computed from the eigen-decomposition of  $\hat{R}(\mathbf{q})$ , and  $L$  is the  $N$ -element vector of ones. Detailed discussion on computational complexity, estimation of the number of sources, and algorithm performances in terms of accuracy and stability for both methods can be found in references [1, 2, 3].

### 3.3 Real-time Implementation Issues

The computational complexity of MUSIC is governed mainly by the eigen-analysis calculation using singular value decomposition (SVD), which is of order  $N^3$  ( $O(N^3)$ ). Other calculations such as matrix-matrix multiplication and matrix inversion are  $O(N^2)$ , or lower [2]. As discussed elsewhere [2], the complexity of the incoherent wideband approach is approximately  $M$  times the number of SVD calculations, and coherent wideband approach is approximately  $S$  times the number of SVD calculations, where  $S$  is the number of look directions or scan angles. For real-time implementation, the size of  $M$  and  $S$  are strictly dependent upon the speed of the signal processing platform.

In the unattended acoustic sensor testbed system, there is no a priori knowledge of the source directions. The sensor arrays scan in all directions (i.e.,  $360^\circ$  field of view). Currently, each sensor array module scans  $360^\circ$  in  $1^\circ$  steps. Therefore, coherent wideband MUSIC requires 360 SVD calculations, which is very computationally intensive for any low-cost, low power single board computer or PC/104 platform with a 486 or Pentium processor. Due to the limitation of the processor, the only way to implement coherent wideband MUSIC in the sensor tested is to apply a fast front-end coarse beamformer to narrow the field of view. We are currently exploring various hybrid and hierarchical techniques to incorporate coherent wideband MUSIC for real-time implementation.

As stated before, most ground vehicles of interest have diesel engines that exhibit a pronounced harmonic structure (i.e., sharp spectra) due to the engine rate and, to a lesser extent, the track slaps; vehicles with diesel engines don't exhibit the broadband energy that the turbine-engine vehicles exhibit [2,3]. The incoherent wideband MUSIC algorithm is, therefore, somewhat more suited for the sensor testbed because of the nature of the acoustic sources and computational complexity. The acoustic sources can be characterized by a set of frequency components; however, the frequency components must have relatively high SNRs. Since the complexity of incoherent wideband MUSIC depends mainly on the number of frequency components  $M$ , real-time implementation is possible for  $M$  as large as 60 depending on the processing power and speed of the CPU. Pre-processing to select the frequency components for wideband processing is very important. Currently, we are using a simple adaptive threshold peak picking algorithm to select the strong frequency components, and the harmonic line association (HLA) algorithm to associate the frequency components based on the engine and track harmonics, and to weed out non-related frequency components [4].

## 4. RESULTS

Table 1 shows the average processing time per update for incoherent wideband MUSIC for  $M=10$  and 20 and three different PC platforms with at least 16 MB RAM and 1 GB hard drive. The results were calculated based on experimental data stored on the respective PC and the average processing time in seconds was based on a 6-minute test run. For each update,  $M=10$  and 20 frequency components with the largest SNR levels were selected for wideband processing. If a dedicated DSP chip can be incorporated into the sensor testbed, at the sensor array level,

to perform the DF algorithm,  $M$  can be as large as 60. Because of the low power requirement, use of an additional processor might not be feasible. Figure 3 shows the pseudospectrogram (a 3-dimensional plot of the beampattern as a function of time) of the incoherent wideband MUSIC implemented in the sensor testbed for a maximum of  $M=10$ . Up to a maximum number of  $M$  frequency components are used for wideband processing that corresponds to the SNR level relative to the set threshold. For this test run of 350 s, there are two identical tanks traveling at the constant speed of 20 mph (except during turns) with 200 m separation. Figure 4 shows the raw DOA estimates of the two targets extracted from the pseudospectrogram in figure 3 versus ground truths. Survey stakes placed on the tank course provided rough location estimates of the ground vehicles as they maneuvered around the course. Figure 5 shows the raw DOA estimates for both delay-sum (figure 5 (a)) and incoherent wideband MUSIC (figure 5 (b)) DF algorithms for the same test run as shown in figure 4, for the time interval [100, 200] s. Closer inspection of the plots shows that the wideband algorithm does a better job of separating the two targets than the delay-sum algorithm, especially near [120, 130] s and [180, 190] s. However, neither algorithm detected both targets for the entire test. In fact, for this particular test run, both algorithms perform DF on both targets only about 40 percent of the time. This might be due to the peak-picking and HLA algorithms selecting the incorrect frequency components (or not selecting any frequency components emanating from the acoustic sources), or one target masking out the other target because of aspect angles. Also it is possible that the terrain and atmosphere absorb and attenuate the signature; therefore, limiting their SNR. Only qualitative analysis is presented because accurate GSP ground truths were not available for these two-target data sets; quantitative analysis such as mean square errors (MSE) cannot be calculated without reliable ground truths.

$M$	66 MHz 486	133 MHz Pentium	200 MHz Pentium Pro
10	4.36 s	0.544 s	0.267 s
20	7.78 s	0.956 s	0.422 s

Table 1. Average processing times per update for wideband incoherent MUSIC with  $M=10$  and 20 for three different PC platforms

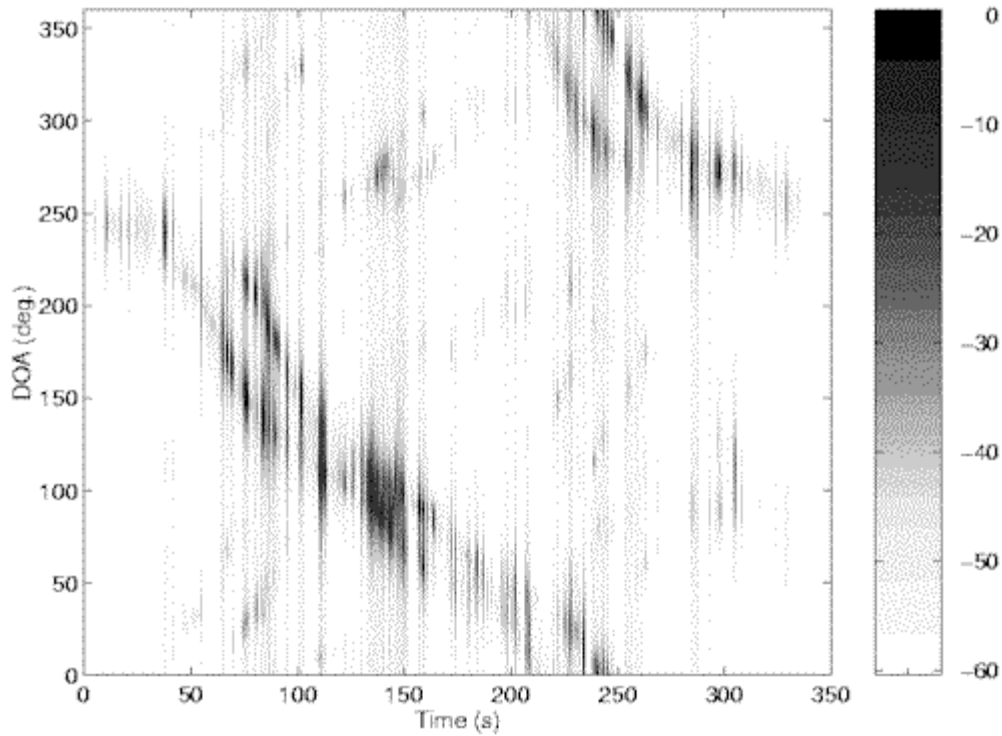


Figure 3. Pseudospectrogram of incoherent wideband MUSIC for two identical tank targets with 200 m separation.

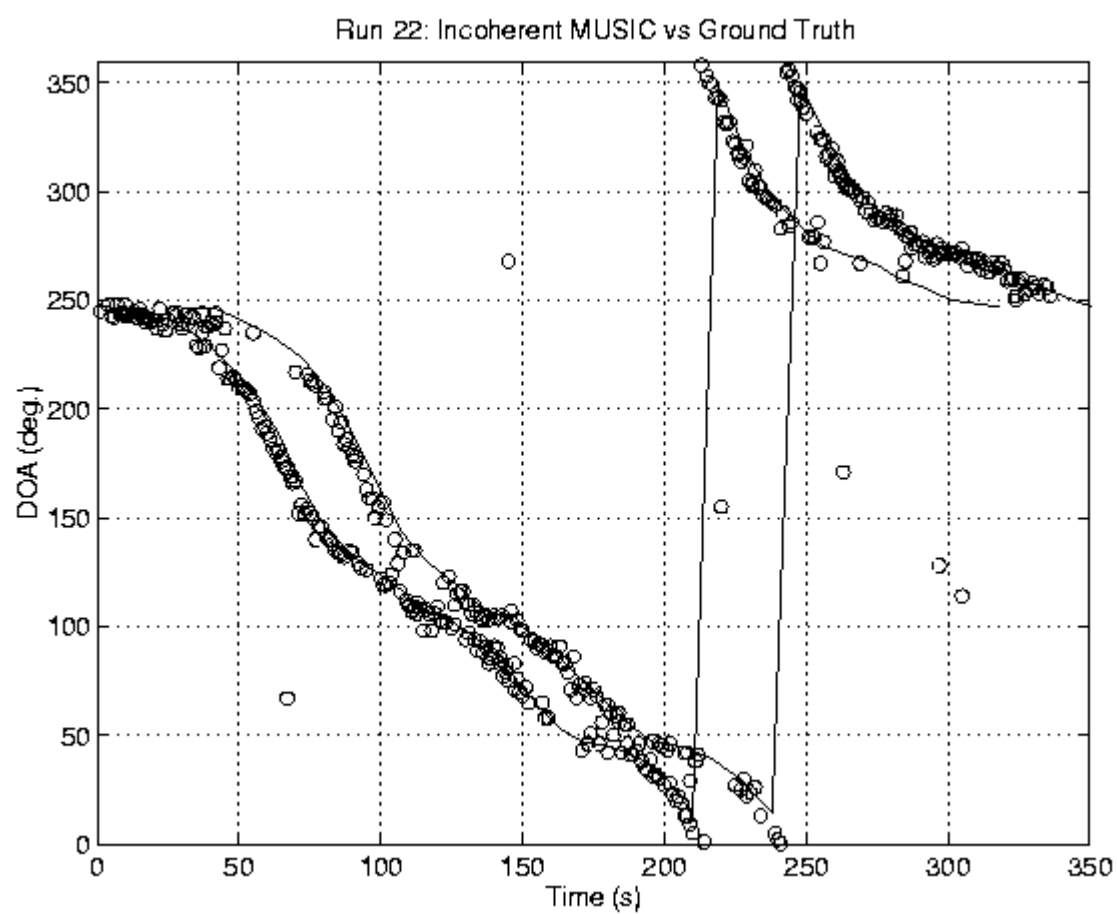


Figure 4. The raw DOA estimates (dots) of incoherent wideband MUSIC versus ground truth (lines) for two identical tank targets with 200 m separation.

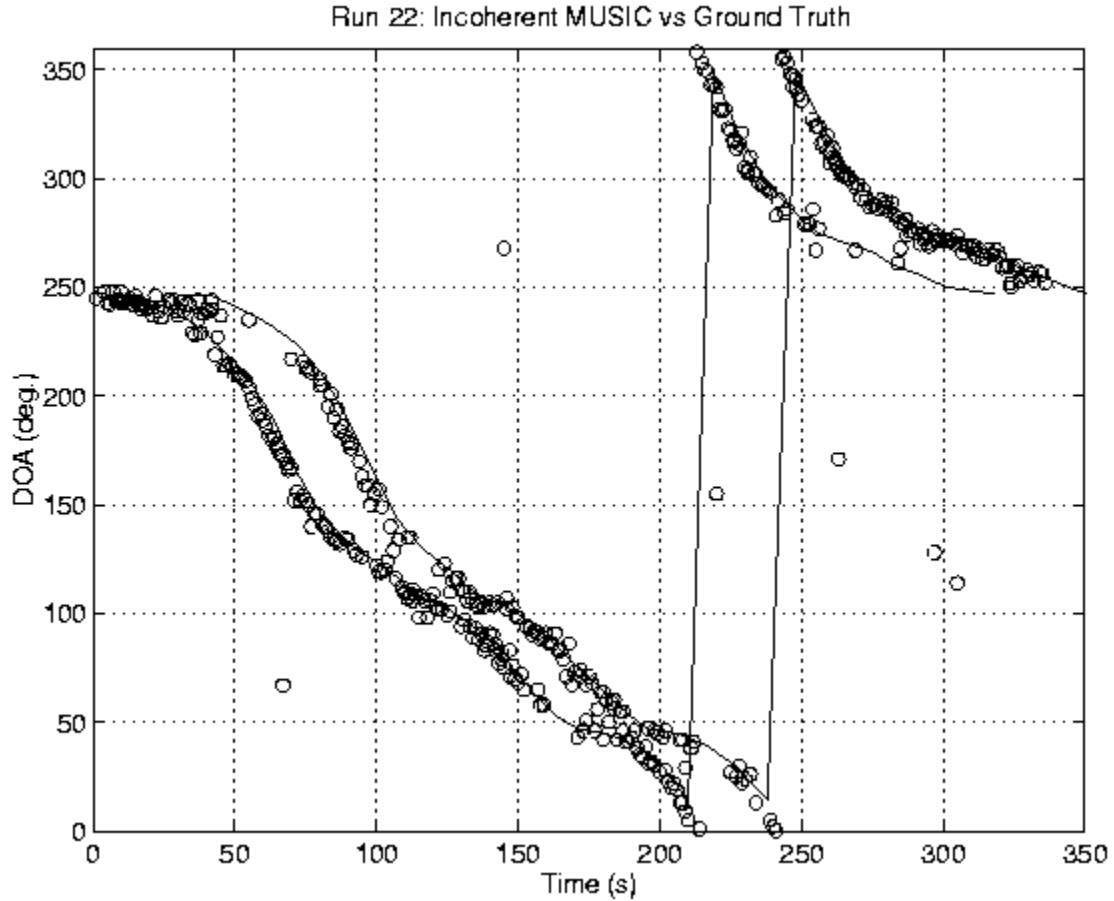


Figure 5. Raw DOA estimates (dots) versus ground truth (lines) for the same test run as in figure 4, for the time interval [100, 200] s, for (a) delay-sum and (b) incoherent wideband MUSIC

## 5. CONCLUSIONS

In this paper, two wideband processing approaches using MUSIC are presented. Real-time implementation issues are discussed with respect to real-time implementation in the ARL acoustic sensor testbed. Based on the current processing hardware used in the sensor testbed, only the incoherent wideband MUSIC can be implemented in real-time with up to  $M \approx 20$  frequency components. However, it will be possible in the very near future to implement both wideband MUSIC algorithms over all frequencies and in all directions of interest with low cost off-the-shelf systems due to the continual advancement in CPU and DSP technology.

Experimental results and analyses for multiple targets were presented using incoherent wideband MUSIC. The acoustic sources are, generally, characterized as a sum of narrowband frequency components. Given adequate SNR, the incoherent wideband MUSIC algorithm performed very well, produced accurate DOA estimates, and yielded sharp distinct peaks in the beampattern. The disadvantage of implementing incoherent instead of coherent wideband MUSIC is stability of the DOA estimates. The coherent approach is more stable than the incoherent approach because it is less sensitive to noise. However, coherent processing requires inclusion of more frequency components to achieve the same accuracy as incoherent processing [2], which will require more calculation. Future work of interest includes developing more robust frequency peak picking and data association algorithms, developing advanced robust tracking algorithms, and incorporating more processing power into the sensor testbed with commercial off-the-shelf hardware.

## ACKNOWLEDGEMENTS

This work was done in-house by the Sensors and Electron Devices Directorate's Acoustic Signal Processing Branch at ARL in Adelphi, Maryland. The authors wish to acknowledge members of the acoustics team and other ARL

engineers for their valuable contributions: J. Eicke, N. Srour, J. Gerber, H. Vu, A. Nguyen, K. Tran, L. Sim, M. Wellman and B. Sadler.

## REFERENCES

1. T. Pham, B. Sadler, M. Fong, and D. Messer, High-resolution direction finding algorithm to detect and track *20<sup>th</sup> Army Science Conference Proceedings*, Vol. 1, pp. 235239 (June 1996).
2. T. Pham and B. Sadler, Adaptive wideband aeroacoustic array processing, *8th IEEE Statistic Signal and Array Processing Workshop (SSAP96) Proceedings*, pp. 295298 (June 1996).
3. T. Pham and B. Sadler, Wideband acoustic array processing to detect and track ground vehicles, *US ARL 1997 Sensors and Electron Devices Symposium Proceedings*, pp. 151454 (January 1997).
4. S. Choy, M. Fong and H. Vu, "Software architecture for an expandable testbed sensor tracking system," ARL internal report (July 1995).
5. N. Srour and J. Robertson, Remote netted acoustic detection system: Final Report, "ARL Technical Report, ARL-TR-706 (April 1995).
6. M. Wellman, N. Srour, and D. Hillis, Acoustic feature extraction for a neural network classifier, "ARL Technical Report ARL-TR-1166 (January 1997).
7. M. Wax, T. Shan, and T. Kailath, Spatio-temporal spectral analysis by eigenstructure methods, *IEEE Trans. ASSAP*, Vol. 32, pp. 817827 (August 1994).
8. H. Wang and M. Kaveh, Coherent signal-subspace processing for the detection and estimation of angles of *IEEE Trans. ASSP*, Vol. 33, pp. 823831 (August 1985).